# Review of Linear Algebra

Chong Jin Ong

Department of Mechanical Engineering,
National University of Singapore

# Outline

## 2.1 Notations and Preliminaries

- $\mathbb{R}^n$ : The $n$-dimensional Euclidean space; $x \in \mathbb{R}^n$ refers to a $n$-dimensional vector of real numbers.
- $A \in \mathbb{R}^{n \times m}$ refers to a $n \times m$ matrix of real numbers.
- Suppose $A \in \mathbb{R}^{n \times m}$, $B \in \mathbb{R}^{m \times r}$, $C \in \mathbb{R}^{\ell \times n}$ and $D \in \mathbb{R}^{r \times p}$. Let $a_i$ be the $i^{th}$ column of $A$ and $b_j$ is the $j^{th}$ row of $B$. Then

$$CA = C \begin{pmatrix} a_1 & a_2 & \cdots & a_m \end{pmatrix} = \begin{pmatrix} Ca_1 & Ca_2 & \cdots & Ca_m \end{pmatrix}$$

$$BD = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} D = \begin{pmatrix} b_1 D \\ b_2 D \\ \vdots \\ b_m D \end{pmatrix}$$

# 2.2 Basis, Representation and Orthonormalization

- **Definition** Linear Independence of vectors: A set of vectors $x_1, x_2, \cdots, x_m \in \mathbb{R}^n$ is said to be linearly dependent if and only if there exists scalars $c_1, c_2, \cdots, c_m$ not all zeros, such that

$$c_1 x_1 + c_2 x_2 + \cdots + c_m x_m = 0$$

  If the only set of $c_i$ such that the above holds is $c_1 = c_2 = \cdots = c_m = 0$, then the set of vectors $x_1, x_2, \cdots, x_m \in \mathbb{R}^n$ is said to be linearly independent.
- The dimension of a linear space is the number of linearly independent vectors in the space. In $\mathbb{R}^n$, we can only find at most $n$ linearly independent vectors.
- **Definition:** A set of linearly independent vectors in $\mathbb{R}^n$ is called a basis if every vector in $\mathbb{R}^n$ can be expressed as a unique linear combination of set.
- In $\mathbb{R}^n$, any set of $n$ linearly independent vectors can be used as a basis.

# Basis, Representation and Orthonomalization

- Let $Q = \{q_1, q_2 \cdots q_n\}$ be a set of l.i. vectors. Then every vector $x$ can be expressed uniquely as

$$x = \alpha_1 q_1 + \alpha_2 q_2 + \cdots + \alpha_n q_n = Q\bar{x}$$

  where $\bar{x}^T = [\alpha_1 \; \alpha_2 \; \cdots \; \alpha_n]$
- $\bar{x}$ is also known as the representation of vector $x$ with respect to basis $Q$.
- For every $\mathbb{R}^n$, there exists the following orthonormal basis

$$i_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad i_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots \quad i_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

  Hence, $I_n = [i_1 \; i_2 \; \cdots i_n]$ is the $n \times n$ unit matrix and any vector $x$ has a representation that is equal to itself with respect to $I_n$.
- Suppose a vector has a representation of $\bar{x}$ in basis $Q$ and a representation of $\hat{x}$ in basis $P$. How are $\bar{x}$ and $\hat{x}$ related?

# Norms of vectors

The concept of norm is a generalization of length of magnitude. Any real-valued function of $x$, denoted by $\|x\|$ can be a norm if it has the following properties:

- $\|x\| \geq 0$ for all $x$ and $\|x\| = 0$ if and only if $x = 0$.
- $\|\alpha x\| = |\alpha|\|x\|$ for all real value $\alpha$.
- $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$ for all $x_1$ and $x_2$ - known as the triangular inequality.

The most commonly used norms are the $\ell_1, \ell_2$ and $\ell_\infty$ norms. For a $x \in \mathbb{R}^n$, they are

- $\|x\|_1 = \sum_{i=1}^n |x_i|$
- $\|x\|_2 = \sqrt{x^T x} = (\sum_{i=1}^n x_i^2)^{0.5}$
- $\|x\|_\infty = max_i |x_i|$

  These are special cases of the $p$-norm, $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{\frac{1}{p}}$

# Norms of Matrices

- Extensions of norms of vectors - put $A \in \mathbb{R}^{n \times m}$ as a big vectors of $nm$ elements.
- A more useful norm is that induced through norm of vectors - induced norms.
- The induced norm of $A$ is the smallest real number $C$ such that

$$\|Ax\| \leq C\|x\|$$

for all $x \in \mathbb{R}^n$. Another way of looking at this is

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\|$$

where sup refers to the supremum or the least upper bound.
- Matrix norm is a measure of the maximum amplification factor brought about by the matrix.
- Since there are $\ell_1, \ell_2$ and $\ell_\infty$ vector norms, they induce corresponding matrix norms.

Norms of matrices also has the following properties

- $\|Ax\| \leq \|A\|\|x\|$
- $\|A+B\| \leq \|A\|+\|B\|$
- $\|AB\| \leq \|A\|\|B\|$

Orthonormal set of vectors

- A vector is said to be normalized if $\|x\|_2 = 1$
- Two vectors $x, y \in \mathbb{R}^n$ are orthogonal if $x^T y = 0$.
- A set of vectors, $x_1, x_2, \cdots, x_m$ is said to be orthonormal if

$$x_i^T x_j = \begin{cases} 0, & \text{if } i \neq j; \\ 1, & \text{if } i = j. \end{cases}$$

# 2.3 Linear Algebraic Equations

Consider the set of linear algebraic equation:

$$Ax = y$$

where $A \in \mathbb{R}^{m \times n}$, $y \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$.

**Domain and Range of a Matrix**

- The matrix $A \in \mathbb{R}^{m \times n}$ has $\mathbb{R}^n$ as its domain. The range of $A$, $\mathcal{R}(A)$, is given by

$$\begin{aligned}\mathcal{R}(A) &= \{y \in \mathbb{R}^m, \text{for which there exists at least one } x \in \mathbb{R}^n \text{ s.t. } y = Ax.\} \\ &= \text{the set of all possible linear combinations of columns of } A\end{aligned}$$

- The dimension of the range space $\mathcal{R}(A)$ is the maximum number of linearly independent columns of $A$.

# 2.3 Linear Algebraic Equations

## Null Space and Nullity of a Matrix

- The null space of matrix $A$ is

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n : \text{ such that } Ax = 0\}$$

- Nullity of $A$ is the number of linearly independent vectors of $\mathcal{N}(A)$ and is denoted by $\nu(A)$.
- Null space of $A$ consists of all its null vectors.
  Remark: If $\nu(A) = 0$, it means that 0 is the only element in $\mathcal{N}(A)$.

## Rank of a Matrix

- The rank of $A \in \mathbb{R}^{m \times n}$, $\rho(A)$, is the maximum number of linearly independent columns or rows in $A$. Hence,

$$\rho(A) \leq \min\{m, n\}$$

  Remarks:
  (i) If $\rho(A)$ equals the number of columns (rows) then $A$ is known as full column (row) rank.
  (ii) If $A$ is square and full rank, then $A$ is non-singular.

## 2.3 Linear Algebraic Equations

**Properties of rank:**

- Let $A \in \mathbb{R}^{m\times n}$ Then $\rho(A) + \nu(A) = n$.
- Let $A \in \mathbb{R}^{q\times n}$ and $B \in \mathbb{R}^{n\times p}$. Then

$$\rho(A) + \rho(B) - n \leq \rho(AB) \leq \min\{\rho(A), \rho(B)\}$$

- Let $A \in \mathbb{R}^{m\times n}$. Then

$$\rho(AC) = \rho(A) \text{ and } \rho(DA) = \rho(A)$$

  for any $n \times n$ and $m \times m$ non-singular matrices $C$ and $D$.
- Given $A \in \mathbb{R}^{m\times n}$, $y \in \mathbb{R}^m$ and $y = Ax$. There exists a vector $x \in \mathbb{R}^n$ satisfying the above equation if and only if $y \in \mathcal{R}(A)$ or equivalently,

$$\rho(A) = \rho([A \quad y])$$

- Given $A \in \mathbb{R}^{m\times n}$. For every $y \in \mathbb{R}^m$, there exists a vector $x \in \mathbb{R}^n$ such that $y = Ax$ if and only if $\rho(A) = m$.

# 2.3 Linear Algebraic Equations

**Determinant of a square matrix:**

- Determinant is a scalar-valued function of a square matrix $A$.
- Can be evaluated via Laplace Expansion:

$$det(A) = |A| = \sum_{j=1}^{n} a_{ij} c_{ij} = \sum_{i=1}^{n} a_{ij} c_{ij}$$

  where $c_{ij}$ is the co-factor corresponding to $a_{ij}$ and

$$c_{ij} = (-1)^{i+j} det(M_{ij})$$

  where $M_{ij}$ is the $(n-1) \times (n-1)$ submatrix of $A$ by deleting the $i^{th}$ row and $j^{th}$ column.
- The determinant of any $r \times r$ submatrix of $A$ is called a minor of order $r$.
- The rank of $A$ is also defined as the largest order of all non-zero minors of $A$.

## 2.3 Linear Algebraic Equations

**Inverse of a square matrix :**

- A square matrix $A$ has an inverse, $A^{-1}$, if and only if $|A| \neq 0$.
- One formula for $A^{-1}$ is based on the co-factor of $A$.
- Let $adj(A)$ be the matrix with the $(i, j)$ element being $c_{ji}$, i.e., $adj(A)$ is the transpose of the matrix of co-factors. Then,

$$A^{-1} = \frac{adj(A)}{det(A)}$$

**Properties of Inverse and Determinant:**

- If any two rows or columns of $A$ are linearly dependent, then $det(A) = 0$.
- $det(A) = det(A^T)$.
- $det(AB) = det(A)det(B)$ if $A$ and $B$ are both square matrices.
- $det(A) = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$ where $\lambda_i$s are the eigenvalues of $A$.
- If $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{m\times n}$ and $D \in \mathbb{R}^{m\times m}$, then

$$det \begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = det \begin{pmatrix} A & 0 \\ C & D \end{pmatrix} = det(A)det(D)$$

## 2.4 Similarity Transformation

- Consider the mapping of $x \in \mathbb{R}^n$ to $y \in \mathbb{R}^n$ in the form of

$$y = Ax$$

  What happens to matrix $A$ when $x$ is represented by a different basis $Q$?

- We show that there exists a matrix $\bar{A}$ such that

$$\bar{y} = \bar{A}\bar{x}$$

  where $\bar{x}$ and $\bar{y}$ are representations of $x$ and $y$ under $Q$.

- Since $x = Q\bar{x}$ and $y = Q\bar{y}$ then

$$y = Ax \Leftrightarrow Q\bar{y} = AQ\bar{x} \Leftrightarrow \bar{y} = Q^{-1}AQ\bar{x}$$

  Hence,

$$\bar{A} = Q^{-1}AQ$$

  is the representation of $A$ in basis $Q$.

- The above expression can also be written as $Q\bar{A} = AQ$, or

$$[q_1\ q_2 \cdots q_n]\bar{A} = A[q_1\ q_2 \cdots q_n] = [Aq_1\ Aq_2 \cdots Aq_n]$$

- In this form, column $i$ of $\bar{A}$ is the representation of $Aq_i$ in basis $Q$.
- $A$ and $\bar{A}$ are said to be similar and the transformation from one to the other is known as similarity transformation.

# 2.5 Diagonal and Jordan Form

As shown earlier, $A$ can have a different representation w.r.t. different bases. Are there bases that are more insightful?

**Definition:** A scalar, $\lambda$, (real or complex) is called an eigenvalue of $A$ if there exists a non-zero vector $x$ (real or complex) such that $Ax = \lambda x$. The vector $x$ is called an (right) eigenvector of $A$ associated with $\lambda$.

- From $Ax = \lambda x, \Leftrightarrow Ax - \lambda Ix = 0 \Leftrightarrow (A - \lambda I)x = 0$
  The above corresponds to $n$ equations with $n$ unknowns.
- For $x$ to be non-zero, $(A - \lambda I)$ must not have full rank.
- Solve for values of $\lambda$ for which $(A - \lambda I)$ loses rank via

$$det(\lambda I - A) = 0$$

**Definition:** The determinant $det(\lambda I - A)$ is called the characteristic polynomial of $A$. It is an $n^{th}$ degree monic polynomial in $\lambda$, which when expanded, yields the characteristic equation

$$det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n = 0$$

# 2.5 Diagonal and Jordan Form

- The $n$ roots of the characteristic equations are known as the eigenvalues of $A$.
- The eigenvector for each eigenvalue $\lambda$ can be obtained from the expression $(A - \lambda I)x = 0$.
- Eigenvectors are unique up to a non-zero scalar multiple.
  **Distinct Eigenvalue**
- Suppose $\lambda_i, i = 1, \cdots, n$ are all distinct with corresponding eigenvectors $q_i$.
- It can be shown that $Q = [q_1\ q_2 \cdots q_n]$ forms a set of linearly independent vectors.
- What is the representation of $A$ under $Q$?

# 2.5 Diagonal and Jordan Form

- Recall that under similarity transformation

$$[q_1 \; q_2 \; \cdots q_n]\bar{A} = [Aq_1 \; Aq_2 \; \cdots Aq_n] = [\lambda_1 q_1 \; \lambda_2 q_2 \; \cdots \lambda_n q_n]$$

- Hence, looking at the first column on both sides, the first column of $\bar{A}$ is

$$[\lambda_1 \; 0 \; 0 \cdots 0]^T$$

  Extending this to the rest of the columns of $\bar{A}$, we have

$$\bar{A} = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

- Conclusion: Every matrix that has distinct eigenvalues can be represented as a diagonal matrix using its eigenvectors as the basis.

# 2.5 Diagonal and Jordan Form

Example: Find the evalues and evectors of

$$A = \begin{pmatrix} 3 & -2 \\ -1 & 4 \end{pmatrix}$$

$$\begin{aligned} det(A - \lambda I) &= \begin{vmatrix} 3-\lambda & -2 \\ -1 & 4-\lambda \end{vmatrix} \\ &= (3-\lambda)(4-\lambda) - 2 = (\lambda - 2)(\lambda - 5) = 0 \end{aligned}$$

For $\lambda = 2:$ $\quad (\lambda I - A)x = \begin{pmatrix} -1 & 2 \\ 1 & -2 \end{pmatrix} x = 0 \quad \Rightarrow x_1 = [2 \quad 1]^T.$

For $\lambda = 5:$ $\quad (\lambda I - A)x = \begin{pmatrix} -2 & -2 \\ -1 & -1 \end{pmatrix} x = 0 \quad \Rightarrow x_1 = [1 \quad -1]^T.$

# 2.5 Diagonal and Jordan Form

**Not all distinct eigenvalues**

- An eigenvalue that has a multiplicity of 2 or higher is known as a repeated eigenvalue.
- Example: $(\lambda - \lambda_1)^2(\lambda - \lambda_2) = 0 \Rightarrow \lambda_1$ has a multiplicity of 2.
- Consider an eigenvalue $\lambda_j$ with multiciplity $m_j$:
  Two cases can happen:

1. If $\nu(A - \lambda_j I) = m_j$, then can find $m_j$ l.i. e-vectors associated with $\lambda_j$
2. If $\nu(A - \lambda_j I) < m_j$, then not possible to find $m_j$ l.i. e-vectors.

- Case 1 is no different from the case of distinct e-values.
- Case 2 means the matrix cannot be diagonalized but can be block diagonalized, known as Jordan Form.
- Needs the concept of generalized e-vectors.

# 2.5 Diagonal and Jordan Form

**Generalized eigenvector (Optional)**

- An vector $v$ is a generalized e-vector of grade $m$ if

$$(A - \lambda I)^m v = 0$$
$$(A - \lambda I)^{m-1} v \neq 0$$

  The standard e-vector correspond to the special case of $m = 1$.

- We illustrate the idea using an example (no intention to develop the theory here!): Suppose $n$ and $\lambda$ is the only repeated e-value of $A$. Assume that $(A - \lambda I)$ has rank 3 and nullity 1. This means that there is only 1 l.i. e-vector $v$. We need 3 more. Assuming that we have $v_2, v_3$ and $v_4$ are generalized e-vectors of grades 2, 3 and 4 respectively and that nullities of $(A - \lambda I)^4, (A - \lambda I)^3$ and $(A - \lambda I)^2$ are all ones. Then, let

$$\begin{aligned} v_4 &:= v \\ v_3 &:= (A - \lambda I)v_4 \\ v_2 &:= (A - \lambda I)v_3 \\ v_1 &:= (A - \lambda I)v_2 \end{aligned} \tag{1a}$$

# 2.5 Diagonal and Jordan Form

**Generalized eigenvector (Optional)**

- Then, it follows from (1a) that $(A-\lambda I)v_1 = (A-\lambda I)^2 v_2 = (A-\lambda I)^3 v_3 = (A-\lambda I)^4 v_4 = 0$ since $v_4$ is a grade 4 eigenvector. Then, it follows that

$$\begin{aligned} Av_1 &= \lambda v_1 \\ Av_2 &= v_1 + \lambda v_2 \\ Av_3 &= v_2 + \lambda v_3 \\ Av_4 &= v_3 + \lambda v_4 \end{aligned}$$

- Using $Q = [v_1 \ v_2 \ v_3 \ v_4]$ as the basis, the representation of $\bar{A}$ in $Q$ is $Q\bar{A} = AQ$ with

$$[v_1 \ v_2 \ v_3 \ v_4]\bar{A} = [Av_1 \ Av_2 \ Av_3 \ Av_4] = [\lambda v_1 \ \lambda v_2 + v_1 \ \lambda v_3 + v_2 \ \lambda v_4 + v_3]$$

$$\Rightarrow \bar{A} = \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

# 2.5 Diagonal and Jordan Form

**Not all distinct eigenvalues**

- In general, suppose $A$ has one eigenvalue $\lambda_1$ with a multiplicity of 3 and $\lambda_2$ with a multiplicity of 1.
- The Jordan form can take one of the following three forms:

$$\begin{pmatrix} \lambda_1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}, \begin{pmatrix} \lambda_1 & 1 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 \\ 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & \lambda_2 \end{pmatrix}$$

- While useful as a analytical tool, the computation of Jordan form is not numerical stable. We will mention Jordan form just to complete the discussions associated with the diagonal form.

## 2.6 Functions of a Square Matrix

**Polynomials of a square matrix**

**Definition:** Let $A$ be a square matrix. If $k$ is a positive integer, we define

$$A^k := A \cdot A \cdots A (k \text{ times, }) \text{ and}$$

$$A^0 = I$$

- Let $f(\lambda)$ be the polynomial $f(\lambda) = \lambda^3 + 2\lambda^2 - 6$, then

$$f(A) = A^3 + 2A^2 - 6I.$$

- If $A$ is a block diagonal, such as $A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ where $A_1$ and $A_2$ are square matrices of appropriate order. It is easy to verify that

$$A^k = \begin{pmatrix} A_1^k & 0 \\ 0 & A_2^k \end{pmatrix} \text{ and } f(A) = \begin{pmatrix} f(A_1) & 0 \\ 0 & f(A_2) \end{pmatrix}$$

- If $A$ and $\bar{A}$ are similar matrices s.t. $\bar{A} = Q^{-1}AQ$, then

$$A^k = A \cdot A \cdots A = (Q\bar{A}Q^{-1})(Q\bar{A}Q^{-1}) \cdots (Q\bar{A}Q^{-1}) = Q\bar{A}^kQ^{-1}$$

## 2.6 Functions of a Square Matrix

**Caley-Hamilton Theorem :** Suppose $A \in \mathbb{R}^{n\times n}$ and its characteristic equation is

$$det(\lambda I - A) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_n = 0, \text{ then}$$

$$A^n + a_1 A^{n-1} + \cdots + a_n I = 0$$

- Implication: $A^r$ where $r \geq n$ can be expressed as linear combinations of $\{I, A, A^2, \cdots, A^{n-1}\}$.

**Theorem 2.1:** Suppose $f(\lambda)$ is given and $A \in \mathbb{R}^{n\times n}$ matrix with char. polynomial

$$det(\lambda I - A) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{n_i}$$

where $n = \sum_{i=1}^{m} n_i$. Define another $(n-1)$ degree polynomial

$$h(\lambda) := \beta_0 + \beta_1\lambda^1 + \beta_2\lambda^2 + \cdots + \beta_{n-1}\lambda^{n-1}$$

with $n$ unknown coefficients $\beta_i$. These unknowns can be obtained by solving the following set of $n$ equations:

$$\frac{d^k f(\lambda)}{d\lambda^k}\Big|_{\lambda=\lambda_i} = \frac{d^k h(\lambda)}{d\lambda^k}\Big|_{\lambda=\lambda_i} \quad \text{for } k = 0, 1, \cdots, (n_i - 1) \text{ and } i = 1, 2, \cdots, m$$

Then we have

$$f(A) = h(A)$$

and we say that $f(\lambda)$ equals to $h(\lambda)$ on the spectrum of $A$.

## 2.6 Functions of a Square Matrix

- Proof of Theorem 2.1 is omitted.
- Using Theorem 2.1, function of a matrix can be easily defined.
- Let $f(\lambda)$ be any function, not necessary polynomial. Then $f(A)$ can be defined. Let $h(\lambda)$ be a polynomial of degree $(n-1)$ where $n$ is the order of $A$. Solve coefficients of $h(\lambda)$ using Thm 2.1 such that $f(\lambda) = h(\lambda)$ on the spectrum of $A$.

Example: Suppose $f(\lambda) = e^{\lambda}$ and $A = \begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix}$. Choose $h(\lambda) = \beta_0 + \beta_1 \lambda$.
Using Theorem 2.1 with $\lambda_i = -1, -2$ means

$$\begin{aligned} e^{-1} &= \beta_0 - \beta_1 \\ e^{-2} &= \beta_0 - 2\beta_1 \end{aligned}$$

Solving $\Rightarrow \beta_0 = 2e^{-1} - e^{-2}$ and $\beta_1 = e^{-1} - e^{-2}$. Then

$$f(A) = h(A) = \beta_0 I + \beta_1 A = \begin{pmatrix} \beta_0 - \beta_1 & \beta_1 \\ 0 & \beta_0 - 2\beta_1 \end{pmatrix}$$

## 2.7 Quadratic Form, Positive and Non-negative Definiteness

- A square matrix $A$ is symmetric if $A = A^T$.
- The scalar function $x^T A x$ where $x \in \mathbb{R}^n, A = A^T \in \mathbb{R}^{n\times n}$ is called a quadratic form.
- A real symmetric matrix $A$ is said to be positive definite if for all $x \in \mathbb{R}^n, x \neq 0, x^T A x > 0$.
- Similarly, a real symmetric matrix $A$ is said to be positive semi-definite if for all $x \in \mathbb{R}^n, x \neq 0, x^T A x \geq 0$.
- A symmetric matrix $A$ is positive definite if all its leading minors are positive i.e.,

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} > 0 \cdots$$

- A symmetric matrix $A$ is positive definite if and only if its eigenvalues are positive.
- If $D \in \mathbb{R}^{n\times m}$ then $DD^T = A$ is positive definite if and only if $D$ has full rank $n$.